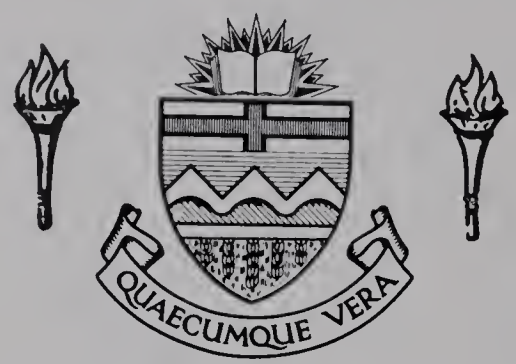


T H E U N I V E R S I T Y O F A L B E R T A

RELEASE FORM

NAME OF AUTHOR ROBIN GARVEN WALKER

TITLE OF THESIS DATA-DEPENDENT TREATMENT ALLOCATION IN TWO-POPULATION SEQUENTIAL TESTING

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Science

YEAR THIS DEGREE GRANTED 1979

THE UNIVERSITY OF ALBERTA

DATA-DEPENDENT TREATMENT ALLOCATION IN TWO-POPULATION SEQUENTIAL TESTING

by

ROBIN GARVEN WALKER

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN

STATISTICS

DEPARTMENT OF MATHEMATICS

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Data Dependent Treatment Allocation in Two-Population Sequential Testing" submitted by Robin Garven Walker in partial fulfilment of the requirements for the degree of Master of Science in Mathematical Statistics.

DEDICATION

To my wife, who wondered if it would ever be done.

ABSTRACT

The clinical comparison of two treatments poses an ethical problem for the experimenter, who wishes to minimize the number of assignments of the poorer treatment while maintaining statistical significance. Sequential methods using data-dependent treatment allocation offer one possible route out of this dilemma. We give a general formulation of two-population sequential testing, and allocation rules for the normal and exponential distributions. Monte Carlo trials indicate that a new rule for the exponential case is superior to previously known rules.

ACKNOWLEDGEMENTS

I wish to thank the following people:

Professor Earl Nordbrock, for suggesting the topic, and his many helpful comments during the early research.

Professor Doug Kelker, for his patience and understanding during the preparation of this thesis.

Mary Willard and Walter Aiello, for their assistance in computer programming.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I Sequential Testing and the Clinical Trial

A. Introduction

One of the most fertile sources of statistical problems has long been the controlled clinical trial. We do not propose to give a detailed history of this field here: this would be a major project in its own right. Rather, we will concentrate our attention upon the use of sequential methods in clinical settings, and some of the problems, both practical and statistical, arising therefrom. In particular, we shall consider some methods of sample-size reduction in two-population hypothesis testing. Sequential estimation is not discussed.

Sequential analysis can be defined as that area of statistics which assumes the sample is drawn unit by unit in time. After the selection of each unit a decision is made whether or not to continue sampling, and if not, to make one of a number of decisions. In its early years, sequential analysis was largely the work of Abraham Wald, whose book (19), remains the classic in the field. The use of sequential methods is suggested by many experimental situations, particularly those in which complete control of the sample is not possible. This is manifestly the case in many medical situations. For example, one cannot usually select, say, fifty candidates for a new cancer treatment: the patients must come to the doctor first, providing a natural basis for sequential analysis.

Medical experimentation has a dimension which it shares with few other fields - an ethical one. Weinstein (20) discusses this problem, and gives a short bibliography. Statistical requirements of sample size often conflict with medical goals. A sequential design

for an experiment is immediately attractive to the medical researcher since it offers the possibility of curtailing administration of an inferior treatment much sooner than would be possible with a fixed-size design. For this reason, medical researchers were early attracted by Wald's methods. The early developments along these lines are well documented in Armitage (3), (4), and Anscombe (2).

The sequential probability ratio test (SPRT) is the centre-piece of Wald's work. However, most clinical researchers have rejected it as unsuitable for medical purposes. To better appreciate this, we now give a brief discussion of this test.

B. The Sequential Probability Ratio Test

Let (Ω, B, P) be a probability space on which is defined a sequence $(X_n)_{n=1}^{\infty}$ of independent, identically distributed random variables which have density $p(x;\theta)$ with respect to P for $\Theta \in H$. To sequentially test the two hypotheses

$$H_o : \Theta = \Theta_o$$
$$H_1 : \Theta = \Theta_1 \ (\neq \Theta_o),$$

two constants a and b are chosen such that $0 < a < 1 < b < \infty$. X_1 is observed, followed by X_2, X_3, and so on. After the n^{th} observation $(n = 1, 2, \ldots)$ the probability ratio L_n is calculated.

$$L_n = \frac{\prod_{i=1}^{n} p(X_i; \Theta_1)}{\prod_{i=1}^{n} p(X_i; \Theta_o)}$$

The following decision rules are used:

1. If $L_n \geq b$, terminate sampling and accept H_1.
2. If $L_n < a$, terminate sampling and accept H_o.
3. If $a \leq L_n \leq b$, observe X_{n+1}.

The constants a and b are chosen such that

$$P \text{ (Accepting } H_1 \mid H_o) \leq \alpha \text{ , and}$$

$$P \text{ (Accepting } H_o \mid H_1) \leq \beta.$$

If H_o is designated as the null hypothesis, then α is the probability of Type I error, β the probability of Type II error. As in the fixed-sample tests, α and β are pre-assigned.

1. Properties of the SPRT

We now state, without proof, the four most important properties of the SPRT. For proofs, we refer the reader to Wald.

1. The test terminates with probability one. That is, a decision is eventually made.
2. $b \leq \frac{1-\beta}{\alpha}$ and $a \geq \frac{\beta}{1-\alpha}$
3. Let $N = \min (n : L_n \notin (a,b))$; that is, N is the size of the sample at termination. Then for all $m < \infty$, $P (N \geq m) > 0$.
4. Among all tests of H_o against H_1 with the same error probabilities, the SPRT minimizes

   $E (N \mid \theta_o)$ and $E (N \mid \theta_1)$.

The fourth property is the test's most medically attractive point. It is the third which renders the SPRT unsuitable for clinical use. No doctor is likely to countenance the possibility of systematically giving a poor treatment to an indefinitely large number of patients.

For this reason much attention has been given to methods of truncating or restricting sequential tests in order to reduce sample sizes. Notable advances in this domain were made by Armitage (3) and Anderson (1).

C. Sample-Size Problem in the Two-Sample Case

In the foregoing discussion we have assumed that there is only one treatment or population in question. When two or more are to be compared, the desiderata for an experimental protocol are somewhat different. In the one-population case, one typically attempts to determine if a treatment is acceptable according to some established criterion. With several competing treatments, the experimenter more often wishes to select the "best", or to rank them. In this setting, the overall sample size becomes less crucial. Instead, it is desired to minimize the number of times the poorer treatments are administered.

In this thesis, we are concerned with the two-sample case. For a comprehensive treatment of k-sample identification and ranking methods, see Bechhofer, Kiefer and Sobel (6), who consider only vector-at-a-time and cyclic sampling, ignoring the sample size problem we are considering. The test we will use is a generalization of the SPRT: its discussion is deferred to the next chapter.

D. The Bechhofer-Kiefer-Sobel Procedure P_B^*

The basic ranking procedure developed by Bechhofer, Kiefer and Sobel for k populations has as its primary goal the selection of the set of the t "best", where t could be anything from 1 to k-1. We shall assume t = 1. Before any observations are taken it is assumed that

each treatment is equally likely to be chosen. After each vector of observations has been made, a posterior probabilities of correct selection are calculated. The first time one of these probabilities exceeds the desired probability of correct selection, the procedure terminates and the treatment corresponding to this maximum probability is selected.

E. Types of Response and Related Probability Models

Medical testing situations fall roughly into three broad categories, each characterized by what we shall refer to as a "response type" and an associated archetypical distribution. Each type has attracted considerable attention as a setting for sequential trials.

1. General Continuous Response

In many cases, patient response to a treatment is such that it allows the assumption of continuity of the response variable. Two examples are:

a) Blood pressure change following administration of a hypotensive agent, and

b) Weight change following a fixed period of supervised diet.

The Normal distribution enters naturally at this point, if for no other reason than the Central Limit Theorem. Many authors have considered sequential tests for normal parameters, and discussion of this area is beyond the scope of this thesis. We will, however, examine some recent results pertaining to the sample-size problem.

2. Discrete Response

In the case of responses falling naturally into a number of

categories, a discrete probability model is appropriate. The most important case by far is that of dichotomous response, with the binomial distribution the related model. Examples of dichotomous response occur wherever treatment results in either success or failure, clearly defined.

Sequential procedures for selecting the best of a set of binomial populations have had much attention in the last decade. One sampling rule that effectively reduces inferior treatment sample sizes is the so-called "Play-the-Winner" rule (PW). Briefly stated, PW sampling in the two-sample case follows each success by a repeat of the treatment, and each failure by a change of treatments. This rule was pioneered by Zelen (21), and Sobel and Weiss (17). Hoel and Sobel (13) compared PW sampling with several other rules, with especially favorable results when both success probabilities are high.

3. Survival Response

In terminal illnesses such as cancer or heart disease, the desired response to treatment is prolongation of life, usually within an observable time-span, such as three to five years. Of the many life distributions, the simplest and most important is the exponential. There is some empirical justification for the use of this distribution as a model for survival in terminal cancer (Zelen (22)), although the presence of non-terminal cases in the sample may invalidate the model to some degree (Boag, discussion of Armitage (3)).

The feature that distinguishes survival response from the general continuous case is the natural censoring of the data. Full information on any patient consists of the time-span from treatment

until death. Throughout the course of an experiment, one must "make do" with partial information, in the form of survival times. In this thesis, we are primarily concerned with the sample-size problem in a survival response situation, with the assumption of exponentially distributed survival times.

F. Summary of the Thesis

In the next chapter, we discuss a general formulation for sequential experiments in the two-sample case. Included therein is a discussion of the sequential likelihood ratio test, and a description of treatment-allocation rules.

Chapter Three develops a two-sample test for comparing two normal means, with attention given to results by Robbins and Siegmund (16) and Flehinger and Louis (9). Chapter Four discusses the corresponding test for the scale parameters of two exponential distributions. Several allocation rules are presented, among them one which appears to be superior to the R_γ rules of Flehinger and Louis (10). The comparisons were made by means of Monte Carlo trials, the results of which are contained in the tables ending chapter Four.

The appendix contains the computer programs used in the Monte Carlo trials, with full documentation.

Chapter II Two-Population Sequential Testing

In this chapter, we present a general formulation for sequential testing when there are two populations, and when data-dependent treatment allocation is used. In addition, we will define criteria according to which allocation rules can be composed.

In a sequential clinical trial, the protocol consists of three rules as follows:

(i) Admission rule: This is the identification of patients eligible for admission to the experiment, which is essentially a medical problem. In the sequel, we shall not deal with the rule, except to assume that admission is not dependent upon the course of the experiment.

(ii) Allocation rule: Once a patient has been admitted to the experiment, he must be assigned one or the other treatment. Rules by which this decision are made are the primary topic of this thesis.

(iii) Termination rule: At some point the admissions must be ceased and a decision made on the relative merits of the treatments. Almost all termination rules are related to the SPRT, discussed in Chapter I.

A. The Hypothesis to be Tested

Suppose that we wish to compare treatments "1" and "2", and that patient responses have densities f_1 and f_2. These densities are assumed to belong to a parametric family

$$F = \{ f(x \mid \theta), \theta \in \Gamma, x \in I \subset R \}$$

We assume $T \subset R^1$. Let

$$D = \{ \Delta : \quad \Theta \in T , \Theta + \Delta \in T , \Theta - \Delta \in T \} .$$

Let $f \in F$, $\Theta \in T$, and $\Delta / 2 \in D$. We define

$$f_1 (x) = f (x \mid \Theta + \Delta / 2), \text{ and} \qquad (2.1)$$

$$f_2 (x) = f (x \mid \Theta - \Delta / 2). \qquad (2.2)$$

In a three-decision problem, we wish to choose one of the following hypotheses:

H_o : $\Delta = 0$; i.e., the treatments are equal,

H_1 : $\Delta = \Delta^*$; i.e., treatment 1 is better,

H_2 : $\Delta = - \Delta^*$; i.e., treatment 2 is better. (2.3)

Here $\Delta^* > 0$ is a constant representing a medically significant difference.

Sometimes the hypothesis H_o may not be of interest, and we wish to choose one of H_1 or H_2 only. We shall refer to this latter formulation as the "two-decision" problem.

B. The Sequential Likelihood Ratio Test

The SPRT cannot be applied directly to test the above, due to the composite nature of the hypothesis. We will adopt as our termination rule a generalization of the SPRT, the sequential likelihood ratio test, or SLRT.

1. Two-decision problem

We will first consider the two-decision problem

$H_1 : \Delta = \Delta^*$, vs.

$H_2 : \Delta = - \Delta^*$.

Let $(X_{11}, X_{12}, \ldots, X_{1n_1})$ be the responses of n_1 subjects administered

treatment 1, and let $(X_{21}, X_{22}, \ldots, X_{2n_2})$ be the responses of n_2 subjects administered treatment 2. We assume the responses are completely independent, and thus the joint likelihood function of these variates, given n_1, n_2 and θ, is defined to be

$$L(X; \theta) = \prod_{i=1}^{n_1} f_1(X_{1i} \mid \theta) \prod_{j=1}^{n_2} f_2(X_{2j} \mid \theta)$$

The likelihood ratio for testing H_1 vs H_2 is then given as

$$\Lambda = \frac{\sup_{\Delta \in H_1,\ \theta \in T} L(X; \theta)}{\sup_{\Delta \in H_2,\ \theta \in T} L(X; \theta)}$$

Substituting (2.1) and (2.2) in this gives

$$\Lambda = \frac{\sup_{\Delta \in H_1,\ \theta \in T} \prod_{i=1}^{n_1} f(X_{1i} \mid \theta + \Delta/2) \prod_{j=1}^{n_2} f(X_{2j} \mid \theta - \Delta/2)}{\sup_{\Delta \in H_2\ \theta \in T} \prod_{i=1}^{n_1} f(X_{1i} \mid \theta + \Delta/2) \prod_{j=1}^{n_2} f(X_{2j} \mid \theta - \Delta/2)}$$

$$\Lambda = \frac{\sup_{\theta \in T} \prod_{i=1}^{n_1} f(X_{1i} \mid \theta + \Delta^*/2) \prod_{j=1}^{n_2} f(X_{2j} \mid \theta - \Delta^*/2)}{\sup_{\theta \in T} \prod_{i=1}^{n_1} f(X_{1i} \mid \theta - \Delta^*/2) \prod_{j=1}^{n_2} f(X_{2j} \mid \theta + \Delta^*/2)} \qquad (2.5)$$

Constants A and B are chosen such that $0 < A < 1 < B < \infty$. Sampling continues until the first time $\Lambda \notin (A,B)$, when H_2 is accepted if $\Lambda < A$ and H_1 is accepted if $\Lambda > B$.

2. Three-Decision Problem

With the addition of $H_o : \Delta = 0$, the test statistic becomes (Λ_1, Λ_2), where

$$\Lambda_i = \frac{\underset{\Delta \in H_i,\ \Theta \in T}{\text{Sup}} L(X;\Theta)}{\underset{\Delta \in H_o,\ \Theta \in T}{\text{Sup}} L(X;\Theta)}, \quad i = 1,2$$

The denominator of this expression is

$$\underset{\Theta \in T}{\text{Sup}} \prod_{i=1}^{n_1} f(X_{1i} \mid \Theta) \prod_{j=1}^{n_2} f(X_{2j} \mid \Theta).$$

The numerator of Λ_1 is identical to the numerator of (2.5), while the numerator of Λ_2 is the denominator of (2.5)

Constants A and B are chosen as in the two-decision case, sampling continuing until $\max(\Lambda_1, \Lambda_2) \notin (A,B)$. H_o is accepted if $\max(\Lambda_1, \Lambda_2) < A$. If $\max(\Lambda_1, \Lambda_2) > B$, H_1 is accepted if $\Lambda_1 > \Lambda_2$, and H_2 is accepted if $\Lambda_2 > \Lambda_1$.

A complete discussion of this test may be found in Chapter 5 of Ghosh (12).

It should be noted that the hypotheses H_1 and H_2 could be generalized to

$H_1^* : \Delta \geq \Delta^*$, vs.

$H_2^* : \Delta \leq -\Delta^*$.

However, even in the simplest cases, finding the suprema of the likelihood functions under these hypotheses is very difficult. Furthermore, termination rules become data-dependent. The tests discussed in this thesis, although based upon the simple hypotheses, may be considered approximations to tests of the composite hypotheses.

C. Allocation Rules

In this section, we present the formulation of a general data-dependent allocation rule in discrete time, as given by Louis (15).

Let $X_{11}, X_{12}, \ldots$ and $X_{21}, X_{22}, \ldots$ be two sequences of independent random variables, identically distributed within the sets, but not necessarily between the sets. Assume at time n there are n_1 X_{1i}'s in the sample and n_2 X_{2j}'s, so that $n = n_1 + n_2$. A new sequence of random variables is defined as follows:

Let
$$z_{n+1} = \begin{cases} X_{1,n_1+1}, & \text{if } k_{n+1} = 1 \\ X_{2,n_2+1}, & \text{if } k_{n+1} = 0 \end{cases},$$

where
$$k_i = \begin{cases} 1, \text{ with probability } p_i\ (z_1, \ldots, z_{i-1}, k_1, \ldots, k_{i-1}) \\ 0, \text{ with probability } 1 - p_i\ (z_1, \ldots, z_{i-1}, k_1, \ldots k_{i-1}) \end{cases}$$

$i > 1$

and
$$p_1 = p,$$
$$0 \leq p_i \leq 1, \quad i > 1.$$

The k_i's constitute a sequence of indicator random variables, not necessarily either independent or identically distributed, which specifies the order in which treatments are assigned. The allocation probabilities, the p_i's, may be allowed to depend upon both past observations (z_i's) and past allocations (k_i's). The allocation rule is completely determined by the constant p (usually taken to be 1/2), and the rules for calculating the p_i's.

It is easily seen that this formulation subsumes both strict alternate sampling and completely randomized sampling. Randomization results from setting p_i equal to a constant for all i, while strict alternation is achieved by taking

$$p_i = \begin{cases} 1 - p_{i-1}, & i \geq 2 \\ p, & i = 1 \end{cases}$$

The general allocation rule permits what might be termed "data-dependent randomization". For administrative reasons, this possibility should be excluded, and we therefore require that $p_i = 0$ or 1 for $i > 1$.

As a general consideration, the p_i's are required to depend upon the observations through the likelihood ratio, because, when a test is close to termination, we would like to sample principally according to the "nearest" hypothesis. In other words, if there is strong evidence for either H_1 or H_2, the suspected poorer treatment should be assigned less frequently than the other. If the trend is towards acceptance of H_o, approximate alternate sampling is desired.

D. Performance Criteria for a Sampling Plan

Three surfaces which together describe the relative merits of a protocol may be defined, in order to compose allocation rules. To emphasize the time-dependence of the likelihood ratio, we shall write Λ as $\Lambda(n)$. Define

$$N = \min \{ n : \Lambda(n) \notin (A,B) \}$$

$$N_i = \min \{ n_i : \Lambda(n) \notin (A,B) \} \quad , \quad i = 1,2$$

These are, respectively, the total sample size and the sample sizes on each treatment at termination. Obviously $N = N_1 + N_2$. The average sample number (ASN) is then defined by

$$ASN = E(N).$$

The inferior treatment number (ITN) is defined by

$$ITN = \begin{cases} E\ (N_i), \text{ where } i = \begin{cases} 1, & \Delta < 0 \\ 2, & \Delta > 0 \end{cases} \\ 1/2\ ASN, \quad = 0 \end{cases}$$

Although we are only considering two- and three-decision problems, we observe here that the ASN and ITN are defined independently of the cardinality of the decision space. The third surface to be defined, the operating characteristic (OC), does not have this independence. We define it here for a k-decision problem, with hypotheses $H_1, \ldots, H_k$.

Let $S = \{ (p_1, \ldots, p_k) : \sum_{i=1}^{k} p_i = 1, p_i \geq 0, i = 1, \ldots, k\}$

Then $OC : D \to S$, such that

$$OC_i\ (\Delta) = P_i\ (\Delta) = P\ (\text{Accepting } H_i \mid \Delta \in D\ , i = 1, \ldots, k)$$

For the two-decision problem, this reduces to the power function of the test and its complement, which is

$$P\ (\Lambda\ (N) \geq B \mid \Delta)$$

In the three-decision problem, the OC is determined by the pair

$$P\ (\Lambda_i\ (N) \geq B \mid \Delta\ , i = 1,2)$$

The general problem we are addressing is the selection of allocation rules which minimize the ITN among all allocation and termination rules with equivalent OC's.

The exact calculation of the operating characteristic of a sequential test is rarely possible. In the case of a SLRT with data-dependent allocation, very few advances have been made. Louis (15) considered the problem of testing for a difference in drift between two Brownian motions. Although space does not permit a full development of Louis's results, we state his principal result as follows:

Theorem: The SLRT with data-dependent allocation for testing for a difference in drift between two Brownian motions has an OC function given by Wald's bounds provided allocation depends on the observations at most through the likelihood function.

By virtue of this theorem, if a test can be shown to be approximately equivalent to testing for a difference in drift between two Brownian motions, the OC of the test can be approximated by Wald's bounds.

E. Summary

An experimental protocol for two-population sequential testing is specified by three-decision rules, of which the admission rule is non-statistical. In this thesis, we restrict our attention to protocols using sequential likelihood ratio tests for termination rules, and consider the choice of allocation rules. Allocation rules may be compared by means of three functions: the average sample number, inferior treatment number, and operating characteristic.

In the next two chapters, we present several examples of experimental protocols.

Chapter III Protocols for Selecting the Larger of Two Normal Means

In this chapter we develop the SLRT for selecting the greater of two normal means. Attention is given to allocation rules applicable to two- and three-decision problems.

A. The Two-Decision Problem

Let $X_1, X_2, \ldots, X_{n_1}, \ldots$ and $Y_1, Y_2, \ldots, Y_{n_2}, \ldots$ be two sequences of independent normally distributed random variables with $E(X_i) = \mu_1$ and $E(Y_j) = \mu_2$, all having variance 1. Robbins and Siegmund (16) have studied a termination and allocation rule for sequential testing of the hypotheses

$$H_1 : \mu_1 > \mu_2, \text{ vs.}$$
$$H_2 : \mu_2 > \mu_1. \qquad (3.1)$$

(These results were first announced by Flehinger, Louis, Robbins and Singer (11)).

Reparameterizing the observations in terms of

$$\Delta = \mu_1 - \mu_2$$

and $$\theta = (\mu_1 + \mu_2)/2$$

changes the hypotheses (3.1) to

$$H_1 : \Delta > 0, \text{ vs.}$$
$$H_2 : \Delta < 0.$$

1. Termination Rule

We present here an alternative derivation of Robbins' and Siegmund's test, in the general framework of Chapter II. A constant $\Delta^* \geq 0$ is chosen,

and the hypotheses to be tested become

$$H_1 : \Delta = \Delta^*, \text{ vs.}$$
$$H_2 : \Delta = -\Delta^* \qquad (3.2)$$

We now develop the SLRT for testing H_1 vs. H_2. To derive the likelihood ratio, we first find the maximum likelihood estimator of θ for Δ fixed. Let $C = (2\pi)^{-n/2}$. Then the likelihood function is

$$L(X,Y;\theta,\Delta) = C \prod_{i=1}^{n_1} \exp[-\tfrac{1}{2}(X_i-\mu_1)^2] \prod_{j=1}^{n_2} \exp[-\tfrac{1}{2}(Y_j-\mu_2)^2]$$

$$= C \prod_{i=1}^{n_1} \exp[-\tfrac{1}{2}(X_i-\theta-\Delta/2)^2] \prod_{j=1}^{n_2} \exp[-\tfrac{1}{2}(Y_j-\theta+\Delta/2)^2]$$

Let $\ell = \ell n\ L$

$$= \ell n\ C - \frac{1}{2} \sum_{i=1}^{n_1} (X_i-\theta-\Delta/2)^2 - \frac{1}{2} \sum_{j=1}^{n_2} (Y_j-\theta+\Delta/2)^2$$

Differentiating with respect to θ, and setting the result equal to zero gives

$$\sum_{i=1}^{n_1} (X_i - \theta - \Delta/2) + \sum_{j=1}^{n_2} (Y_j - \theta + \Delta/2) = 0. \qquad (3.3)$$

Let $\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$, and $\bar{Y} = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_j$. Then (3.3) is

$$n_1\bar{X} - n_1\theta - n_1\ \Delta/2 + n_2\bar{Y} - n_2\theta + n_2\ \Delta/2 = 0,$$

whose solution for θ is

$$\hat{\theta} = \frac{n_1\bar{X} + n_2\bar{Y}}{n} - \frac{\Delta}{2}\left(\frac{n_1 - n_2}{n}\right)$$

$$\text{Denote by} \quad \hat{\theta}_+ = \frac{n_1\bar{X} + n_2\bar{Y}}{n} - \frac{\Delta^*}{2}\left(\frac{n_1 - n_2}{n}\right)$$

$$\hat{\theta}_- = \frac{n_1\bar{X} + n_2\bar{Y}}{n} + \frac{\Delta^*}{2}\left(\frac{n_1 - n_2}{n}\right) \qquad (3.4)$$

The likelihood ratio then becomes

$$\Lambda(n) = \frac{L(X,Y;\hat{\theta}_+,\Delta^*)}{L(X,Y;\hat{\theta}_-,-\Delta^*)}$$

$$= \prod_{i=1}^{n_1} \exp\{-\tfrac{1}{2}[(X_i-\hat{\theta}_+-\frac{\Delta^*}{2})^2 - (X_i-\hat{\theta}_-+\frac{\Delta^*}{2})^2]\}$$

$$\cdot \prod_{j=1}^{n_2} \exp\{-\tfrac{1}{2}[(Y_j-\hat{\theta}_++\frac{\Delta^*}{2})^2 - (Y_j-\hat{\theta}_--\frac{\Delta^*}{2})^2]\} \qquad (3.5)$$

Upon substitution for $\hat{\theta}_+$ and $\hat{\theta}_-$, the first product in (3.5) becomes

$$\prod_{i=1}^{n_1} \exp\{(\frac{n_1\bar{X} + n_2\bar{Y}}{n} - X_i)(\Delta^*(\frac{n_1 - n_2}{n}) - \Delta^*)\}$$

$$= \prod_{i=1}^{n_1} \exp\{\frac{-2\Delta^* n_2}{n}\left(\frac{n_1\bar{X} + n_2\bar{Y}}{n} - X_i\right)\}$$

$$= \exp[\frac{-2\Delta^* n_1 n_2}{n}\left(\frac{n_1\bar{X} + n_2\bar{Y}}{n}\right) + 2\Delta^*\frac{n_1 n_2}{n}\bar{X}] \qquad (3.6)$$

Similarly, the second product in (3.5) becomes

$$\exp[\frac{2\Delta^* n_1 n_2}{n}\left(\frac{n_1\bar{X} + n_2\bar{Y}}{n}\right) - 2\Delta^*\frac{n_1 n_2}{n}\bar{Y}] \qquad (3.7)$$

Multiplying (3.6) and (3.7) gives the likelihood ratio

$$\Lambda(n) = \exp[2\Delta^*\frac{n_1 n_2}{n}(\bar{X} - \bar{Y})].$$

The continuation region of the test is given as $\Lambda \in (A,B)$, where

$0 < A < 1 < B < \infty$. If we define

$$a = \frac{\ell n\ A}{2\Delta^*} \quad \text{and} \quad b = \frac{\ell n\ B}{2\Delta^*},$$

the test may be based upon the statistic

$$\zeta(n) = \frac{n_1 n_2}{n} (\overline{X} - \overline{Y}), \tag{3.8}$$

with termination occurring the first time $\zeta(n) \notin (a,b)$. If $\zeta(n) \leq a$, H_2 is accepted. If $\zeta(n) \geq b$, H_1 is accepted.

2. The Operating Characteristic of the Test

The statistics $\zeta(n)$ are linear functions of jointly normal random variables, and are therefore jointly normally distributed for all n. We have

$$E_\Delta \zeta(n) = \frac{n_1 n_2}{n} E(\overline{X} - \overline{Y})$$

$$= \frac{n_1 n_2}{n} \Delta \text{ , and} \tag{3.8}$$

$$Var_\Delta(\zeta(n)) = \frac{n_1^2 n_2^2}{n^2} Var(\overline{X} - \overline{Y})$$

$$= \frac{n_1^2 n_2^2}{n^2} [Var(\overline{X}) + Var(\overline{Y})]$$

$$= \frac{n_1^2 n_2^2}{n^2} (\frac{1}{n_1} + \frac{1}{n_2})$$

$$= \frac{n_1 n_2}{n} .$$

Thus, the distribution of $\zeta(n)$ does not depend on θ, and the test is, in fact, a SPRT for testing H_1 vs. H_2. (This latter fact is perhaps more

clearly seen via the derivation by Robbins and Siegmund (16), who employ a linear transformation of the observations to eliminate θ.)

Therefore, for any allocation rule for which termination occurs with probability 1, Wald's bounds approximate the error probabilities of the test. Thus, we have

$$P_{-\Delta}\ (\text{accept } H_1) < B^{-1}[1 - P_{\Delta}\ (\text{accept } H_2)]$$

and $$P_{\Delta}\ (\text{accept } H_2) < A[1 - P_{-\Delta}\ (\text{accept } H_1)].$$

There is approximate equality in these statements, so simultaneous solution gives

$$P_{-\Delta}\ (\text{accept } H_1) \overset{\sim}{=} \frac{1 - A}{B - A}\ , \text{ and}$$

$$P_{\Delta}\ (\text{accept } H_2) \overset{\sim}{=} \frac{A(B - 1)}{B - A}\ .$$

In terms of a and b, these become

$$P_{-\Delta}\ (\text{accept } H_1) \overset{\sim}{=} \frac{1 - e^{2\Delta^* a}}{e^{2\Delta^* b} - e^{2\Delta^* a}}\ , \text{ and} \tag{3.9}$$

$$P_{\Delta}\ (\text{accept } H_2) \overset{\sim}{=} \frac{e^{2\Delta^* a}(e^{2\Delta^* b} - 1)}{e^{2\Delta^* b} - e^{2\Delta^* a}}\ . \tag{3.10}$$

These errors probabilities do not depend on the particular allocation rule used, provided the rule depends on the observations at most through the differences $X_i - Y_j$.

If the allocation rule is symmetric in X_i and Y_j, and $a = -b$, we have $P_{-\Delta}\ (\text{accept } H_1) = P_{\Delta}\ (\text{accept } H_2)$. Then (3.9) and (3.10) give

$$P_{\Delta}\ (\text{error}) < \frac{1}{1 + e^{-2\Delta a}}\ , \ \Delta \neq 0. \tag{3.11}$$

Again, there is approximate equality.

3. Estimates of Expected Sample Sizes

If $\mathcal{F}_{n_1,n_2}$ denotes the σ-algebra generated by the x_i's and the y_j's, it may be shown that the observed process $\{\zeta(n) - \frac{n_1 n_2}{n}\Delta, \mathcal{F}_{n_1,n_2}\}$ is a martingale. Robbins and Siegmund exploit this fact to estimate expected sample sizes in the symmetric case referred to in the last paragraph of the preceding section.

Their results can be summarized as follows:

1. For $\Delta \neq 0$

$$E_\Delta(\frac{N_1 N_2}{N}) \tilde{=} \frac{b}{\Delta}\,(\frac{e^{2b\Delta} - 1}{e^{2b\Delta} + 1}) \tag{3.12}$$

If $\Delta = 0$

$$E_\Delta(\frac{N_1 N_2}{N}) \tilde{=} b^2$$

2. Strict alternate sampling has $N_1 = N_2 = N/2$, so (3.12) gives

$$E_\Delta(N) = \frac{4b}{\Delta}\,(\frac{e^{2b\Delta} - 1}{e^{2b\Delta} + 1})$$

$$\text{and } E_\Delta(N_1) = E_\Delta(N_2) = \frac{2b}{\Delta}\,(\frac{e^{2b\Delta} - 1}{e^{2b\Delta} + 1}) \tag{3.13}$$

3. Since $\min(N_1,N_2) \geq \frac{N_1 N_2}{N}$, for <u>any</u> allocation rule we have

$$\min[E_\Delta(N_1), E_\Delta(N_2)] \geq E_\Delta\,(\frac{N_1 N_2}{N}) \tag{3.14}$$

4. Simple calculations show that, for all $N_1 + N_2 = N$

$$N \geq \frac{4N_1N_2}{N}$$

Therefore

$$E_\Delta N \geq 4E_\Delta \left(\frac{N_1N_2}{N}\right). \tag{3.15}$$

Equality holds only if $P_\Delta(N_1=N_2) = 1$.

There are two important consequences of these results. First, (3.12) through (3.14) show that the expected sample size on each treatment must exceed approximately one-half the expected number on that treatment under alternate sampling. Second, (3.15) states that the ASN is (approximately) minimized under alternate sampling.

4. Allocation Rule

Robbins and Siegmund studied an allocation for the symmetric test discussed in the preceding sections. We formulate this rule in the general form of Chapter II. A constant $c > 0$ is chosen.

$$\text{Define } I(S) = \begin{cases} 1 \text{ if } S \text{ is true} \\ 0 \text{ if } S \text{ is false} \end{cases}$$

$$\text{Let } Z_{n+1} = \begin{cases} X_{n_1+1}, & \text{if } k_{n+1} = 1 \\ Y_{n_2+1}, & \text{if } k_{n+1} = 0 \end{cases}$$

$$\text{where } k_i = \begin{cases} 1, \text{ with probability } I\left[\frac{n_1-n_2}{n} \leq \frac{\zeta(n)}{c}\right], & i > 2 \\ 0, \text{ with probability } 1 - I\left[\frac{n_1-n_2}{n} \leq \frac{\zeta(n)}{c}\right], & i > 2 \\ 1, & i = 1 \\ 0, & i = 2 \end{cases}$$

That is, we begin by observing X_1 and Y_1. For succeeding observations,

if termination has not occurred we observe an X if

$$\frac{n_1 - n_2}{n} \leq \frac{\zeta(n)}{c} ,$$

otherwise, a Y is observed.

Monte Carlo results on sample sizes and OC's were obtained by Robbins and Siegmund (16). A selection of these results is presented in Table 1 (see Table 1, page 24). The sampling rule appears to come very close to achieving the theoretical bounds given in section A.4.

B. The Three-Decision Problem

We now add the "indifference" hypothesis $H_o : \Delta = 0$ to the problem stated in section A.1. Flehinger and Louis (9) have studied a termination rule and a class of allocation rules for the three-decision case.

1. Termination Rule

We shall use the same nomenclature as in the two-decision problem. Denoting by

$$\hat{\theta}_o = \frac{n_1\bar{X} + n_2\bar{Y}}{n} ,$$

the likelihood ratios are this

$$\Lambda_1(n) = \frac{L(X,Y;\hat{\theta}_+,\Delta^*)}{L(X,Y;\hat{\theta}_o,0)} , \text{ and}$$

$$\Lambda_2(n) = \frac{L(X,Y;\hat{\theta}_-,-\Delta^*)}{L(X,Y;\hat{\theta}_o,0)}$$

TABLE 1

MONTE CARLO RESULTS FOR ROBBINS-SIEGMUND

ALLOCATION RULE: SYMMETRIC TEST WITH b = 6

Δ	OC			ASN			ITN		
	WA[1]	c=6	c=7.2	WA (alternate)	c=6	c=7.2	WA (alternate)	c=6	c=7.2
0	-	-	-	144	216	191	72	108	95
.1	.77	.78	.75	128	197	166	64	86	80
.25	.95	.96	.97	86	146	126	43	34	36
.5	.50	.996	1.00	48	109	78	24	16	16
1.0	1.00	1.00	1.00	24	71.1	45.2	12	7.2	7.7

(1) WA = Wald approximation

$$\Lambda_1(n) = \frac{L(X,Y;\hat{\theta}_+,\Delta^*)}{L(X,Y;\hat{\theta}_o,0)}$$

$$L(X,Y;\hat{\theta}_+,\Delta^*) = C \prod_{i=1}^{n} \exp[-\tfrac{1}{2}(X_i-\hat{\theta}_+-\frac{\Delta^*}{2})^2] \prod_{j=1}^{n_2} \exp[-\tfrac{1}{2}(Y_j-\hat{\theta}_++\frac{\Delta^*}{2})]$$

$$\hat{\theta}_+ + \frac{\Delta^*}{2} = \frac{n_1\bar{X} + n_2\bar{Y}}{n} - \frac{\Delta^*}{2}\left(\frac{n_1 - n_2}{n}\right) + \frac{\Delta^*}{2}$$

$$= \frac{n_1\bar{X} + n_2\bar{Y}}{n} - \frac{\Delta^*}{2}\left(\frac{n_1 - n_2}{n} - 1\right)$$

$$= \frac{n_1\bar{X} + n_2\bar{Y} + n_2\Delta^*}{n}$$

$$\hat{\theta}_+ - \frac{\Delta^*}{2} = \frac{n_1\bar{X} + n_2\bar{Y} - n_1\Delta^*}{n} \quad .$$

$$L(X,Y;\hat{\theta}_o,0) = C \prod_{i=1}^{n_1} \exp[-\tfrac{1}{2}(X_i - (\frac{n_1\bar{X} + n_2\bar{Y}}{n}))^2] \prod_{j=1}^{n_2} \exp[-\tfrac{1}{2}(Y_j - (\frac{n_1\bar{X} + n_2\bar{Y}}{n}))^2]$$

$$\Lambda_1(n) = \prod_{i=1}^{n_1} \exp\{\tfrac{1}{2}([X_i - (\frac{n_1\bar{X} + n_2\bar{Y}}{n})]^2 - [X_i - (\frac{n_1\bar{X} + n_2\bar{Y} + n_2\Delta^*}{n})]^2)\}$$

$$\cdot \prod_{j=1}^{n_2} \{\tfrac{1}{2}([Y_j - (\frac{n_1\bar{X} + n_2\bar{Y}}{n})]^2 - [Y_j - (\frac{n_1\bar{X} + n_2\bar{Y} - n_1\Delta^*}{n})]^2\}$$

The first product equals:

$$\prod_{i=1}^{n_1} \exp\{\tfrac{1}{2}(2X_i - 2(\frac{n_1\bar{X} + n_2\bar{Y}}{n}) - \frac{n_2\Delta^*}{n})(\frac{n_2\Delta^*}{n})\} \qquad (3.16)$$

The second product equals:

$$\prod_{j=1}^{n_2} \exp\{\tfrac{1}{2}(2Y_j - 2(\frac{n_1\bar{X} + n_2\bar{Y}}{n}) + \frac{n_1\Delta^*}{n})(-\frac{n_1\Delta^*}{n})\} \qquad (3.17)$$

(3.16) becomes

$$\exp\{\frac{n_2\Delta^*}{n} \sum_{i=1}^{n_1} [X_i - (\frac{n_1\bar{X} + n_2\bar{Y}}{n}) - \frac{n_2\Delta^*}{2n}]\}$$

$$= \exp\{\frac{n_2\Delta^*}{n} [n_1\bar{X} - \frac{n_1}{n}(n_1\bar{X} + n_2\bar{Y}) - \frac{n_1n_2\Delta^*}{2n}]\}$$

$$= \exp\{\frac{n_2\Delta^*}{n} [\frac{n_1n_2}{n}\bar{X} - \frac{n_1n_2}{n}\bar{Y} - \frac{n_1n_2\Delta^*}{2n}]\}$$

$$= \exp\{\frac{n_1n_2^2}{n^2}\Delta^* [\bar{X} - \bar{Y} - \frac{\Delta^*}{2}]\}$$

(3.17) becomes

$$\exp\{\frac{-n_1\Delta^*}{n} [\sum_{j=1}^{n_2} [Y_j - \frac{n_2}{n}(n_1\bar{X} + n_2\bar{Y}) + \frac{n_1\Delta^*}{2n}]\}$$

$$= \exp\{\frac{-n_1^2n_2\Delta^*}{n^2} [\bar{Y} - \bar{X} + \frac{\Delta^*}{2}]\}$$

Thus,

$$\Lambda_1(n) = \exp\{\frac{n_1n_2}{n}\Delta^*(\bar{X} - \bar{Y} - \frac{\Delta^*}{2})\}$$

Because $\frac{\Lambda_1(n)}{\Lambda_2(n)} = \Lambda(n)$ (from the definition), we have immediately

$$\Lambda_2(n) = \exp\{\frac{n_1n_2}{n}\Delta^*(\bar{Y} - \bar{X} - \frac{\Delta^*}{2})\}$$

As in the two-decision problem, two constants, A and B, are chosen, such that $0 < A < 1 < B < \infty$. The procedure terminates when $\max(\Lambda_1,\Lambda_2)$ leaves the interval (A,B).

2. A Class of Allocation Rules

We now define the rules R_γ. A constant γ is selected with $0 \leq \gamma \leq 1$. Define $\hat{\Delta} = \bar{X}_{n_1} - \bar{Y}_{n_2}$. Let $p_1 = 1/2$, and

$$p_{n+1} = \begin{cases} 1, & n_1 - n_2 \geq \gamma\,(n+1) \text{ or } |n_1 - n_2| < \gamma(n+1) \text{ and } \hat{\Delta} \geq 0 \\ 0, & n_1 - n_2 \leq -\gamma(n+1) \text{ or } |n_1 - n_2| < \gamma(n+1) \text{ and } \hat{\Delta} < 0; \end{cases}$$

where $p_{n+1} = p_{n+1}(Z_1,\ldots,Z_n,k_1,\ldots,k_n)$ is as defined in section 2.C.

When $\gamma = 0$, the rule is strict alternation, and when $\gamma = 1$, it assigns the leading treatment at every arrival.

Flehinger and Louis (9) obtained Monte Carlo results for the R_γ rules, using the stopping rule discussed in the preceding section. Table 2 (see Table 2 on page 28) contains a selection of their results.

TABLE 2

MONTE CARLO RESULTS

NORMAL DISTRIBUTION - THREE-DECISION PROBLEM

WITH A = 0.1, B = 30

Δ	OC		ASN		ITN	
γ	.2	.5	.2	.5	.2	.5
0	.05	.05	127	160	64	80
.125	.13	.14	141	180	63	66
.25	.43	.43	164	211	68	62
.50	.94	.94	107	136	43	36
1.0	1.00	1.00	40	51	16	13

CHAPTER IV PROTOCOLS FOR SELECTING THE BETTER OF TWO EXPONENTIAL PARAMETERS

In this chapter we examine allocation and termination rules for selecting the better of two treatments, when patients' survival times are exponentially distributed. Our attention shall be given entirely to the three-decision problem, and the comparison of several allocation rules.

A. The Testing Situation

We consider a situation in which suitable patients arrive for treatment at a rate of one patient per unit time. There is no loss of generality in this assumption, since time units may be re-defined as necessary. Patients are assigned treatments by an allocation rule, and their survival times after treatment are recorded.

We assume survival time of patients receiving treatment i follows an exponential distribution with instantaneous death rate λ_i. That is, if X is survival time then the density of X is

$$f(x) = \begin{cases} \lambda_i e^{-\lambda_i x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

We will also have occasion to use the parameter $\mu_i = \lambda_i^{-1}$, which is the mean of the distribution.

As in previous chapters, the subjects are assumed to be independent.

A constant $\rho^* > 1$ is chosen which indicates a medically significant difference in lifetime. We wish to test the three hypotheses

$$H_0 : \lambda_1 = \lambda_2$$
$$H_1 : \lambda_2 = \rho^* \lambda_1 \qquad (4.1)$$
$$H_2 : \lambda_1 = \rho^* \lambda_2$$

(These hypotheses can be written in the form of section 2.A, but the formulation above is more natural for the exponential case.)

It is further assumed that data on each patient is updated continuously. This being the case, at each time n, we have the following data:

n_{in}, i = 1,2, the number of subjects assigned treatment i, with $n_{1n} + n_{2n} = n$;

d_{in}, i = 1,2, the number of deceased subjects on treatment i;

X_{ijn}, i = 1,2,; j = 1,2,...,n_{in}, the survival time of the j^{th} patient on treatment i up to time n (4.2)

To distinguish between active patients and those who have died, the X_{ijn}'s are separated into two subsets for computational purposes:

x_{ijn}, i = 1,2; j = 1,...,d_{in}, the survival time after treatment of a deceased patient;

y_{ijn}, i = 1,2; j = 1,...,$n_{in} - d_{in}$, the time since treatment of an active patient (4.3)

The survival times are summarized by the statistics t_{1n} and t_{2n}, where

$$t_{in} = \sum_{j=1}^{n_{in}} X_{ijn}, \quad i = 1,2 \qquad (4.4)$$

That is, t_{in} is the total survival time on treatment i up to time n. In the sequel, we shall suppress the n subscripts for simplicity of notations.

1. Graphical Representation of Data

There is a convenient way of representing this type of data graphically. Figure 1 illustrates this method for a small set of data, 12 observations with $\mu = 10$, and 11 with $\mu = 5$, under alternating treatment allocation (see Figure 1 on page 32).

The sample from $\mu = 10$ gave values 4.3, 4.7, 4.7, etc., and the sample from $\mu = 5$ gave values 4.5, 5.0, 1.4, etc. At time 24, shown on the graph, there are 9 patients surviving from the original 23; 4 of the survivors had received treatment 2, the rest treatment 1.

B. Maximum Likelihood Estimation of the Death Rate

A maximum likelihood estimate of the instantaneous death rate is required in the derivation of the SLRT. The "usual" estimate is not applicable in this situation, since the data are partially censored by survival.

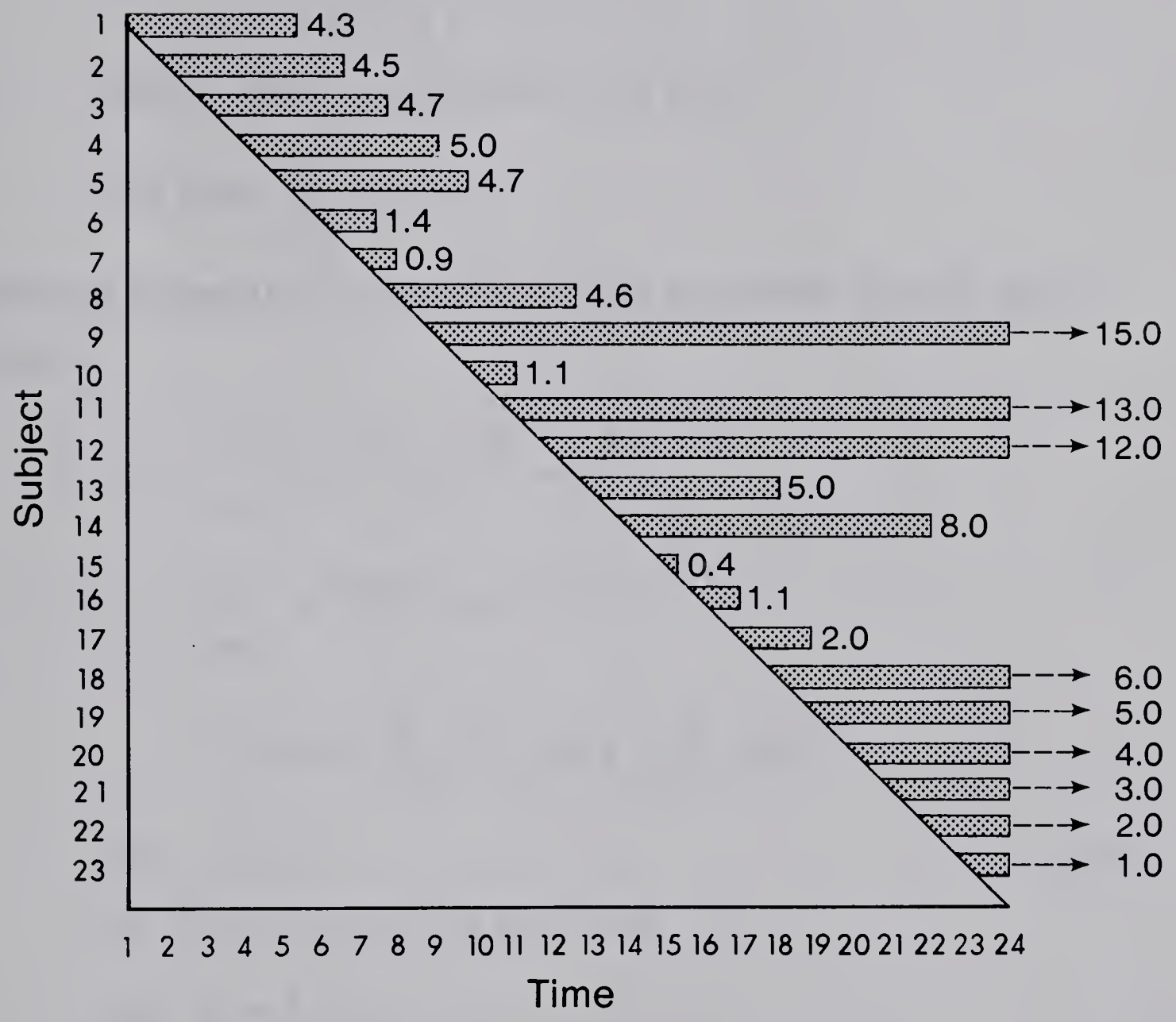
This estimation problem was studied by Bartholomew (5), whose main result is expressed in the following

Theorem: Let X_j, $j = 1,\ldots,n$ be a random sample from an exponential distribution with death rate λ. Define d and t as in (4.2) and (4.4). Then the maximum likelihood estimate of λ is

$$\hat{\lambda} = d/t, \qquad (4.5)$$

provided entry times are independent of previous data.

Graphical Representation of Survivor Data



---→ Surviving Subject

μ_1 (odd numbers) = 10
μ_2 (even numbers) = 5

$n_1 = 12$	$d_1 = 7$	$t_1 = 59.0$
$n_2 = 11$	$d_2 = 7$	$t_2 = 49.7$

Proof: Define P_j = P[subject has died | entry X_j ago], j = 1,...,n

$$Q_j = 1 - P_j = e^{-\lambda X_j}$$

$$\varepsilon_j = \begin{cases} 1, & X_j = x_j, \ j = 1,\ldots,n \\ 0, & X_j = y_j \end{cases}$$

where x_i and y_i are defined as in 4.3.

Note that $\sum_{j=1}^{n} \varepsilon_j = d$.

Under the assumption of independence, the likelihood function may be written:

$$L(\lambda) = \prod_{j=1}^{n} P_j^{\varepsilon_j} Q_j^{1-\varepsilon_j} \{\lambda e^{-\lambda X_j}/P_j\}^{\varepsilon_j}$$

$$= \prod_{j=1}^{n} (e^{-\lambda y_j})^{1-\varepsilon_j} (\lambda e^{-\lambda x_j})^{\varepsilon_j}$$

$$= \lambda^d \exp[-\lambda(\sum_{j=1}^{n} (1 - \varepsilon_j)y_j + \sum_{j=1}^{n} \varepsilon_j x_j)]$$

$$L(\lambda) = \lambda^d e^{-\lambda t} \tag{4.6}$$

Let $\ell(\lambda) = \ell n\, L(\lambda) = d\, \ell n\, \lambda - \lambda t$

Then $\frac{\partial \ell}{\partial \lambda} = \frac{d}{\lambda} - t$

Setting this equal to zero gives the desired result.

C. The SLRT for the Exponential Parameters

Returning to the setting of sections 4.A, we now develop the SLRT for the hypotheses (4.1). Reinstating the i subscripts in (4.6) gives

$$L_i(\lambda_i) = \lambda_i^{d_i} e^{-\lambda_i t_i}, \ i = 1,2$$

Assuming that treatment assignment is independent of past data, the joint likelihood function of λ_1 and λ_2 is

$$L(\lambda_1,\lambda_2) = L_1(\lambda_1)L_2(\lambda_2)$$

$$= \lambda_1^{d_1}\exp(-\lambda_1 t_1)\lambda_2^{d_2}\exp(-\lambda_2 t_2)$$

When $\lambda_2 = \rho\lambda_1$, this becomes

$$L = \rho^{d_2}\lambda_1^{d_1+d_2}\exp[-\lambda_1(t_1 + \rho t_2)]$$

For fixed d_1, d_2, t_1, t_2, this is maximized by

$$\hat{\lambda}_1 = \frac{d_1 + d_2}{t_1 + \rho t_2}$$

Recalling that we are interested in the hypotheses (4.1), we must find the maximum values of $L(\lambda_1,\lambda_2)$ subject to $\rho = 1$, $\rho = \rho^*$, $\rho = 1/\rho^*$, respectively. These maximum values are

$$L_0 = \left(\frac{d_1 + d_2}{t_1 + t_2}\right)^{d_1+d_2}\exp(-d_1 - d_2)$$

$$L_1 = \rho^{*d_2}\left(\frac{d_1 + d_2}{t_1 + \rho^* t_2}\right)^{d_1+d_2}\exp(-d_1 - d_2)$$

$$L_2 = \rho^{*d_1}\left(\frac{d_1 + d_2}{\rho^* t_1 + t_2}\right)^{d_1+d_2}\exp(-d_1 - d_2)$$

The likelihood ratio pair (Λ_1,Λ_2) is thus given by

$$\Lambda_1 = \frac{L_1}{L_0} = \rho^{*d_2}\left(\frac{t_1 + t_2}{t_1 + \rho^* t_2}\right)^{d_1+d_2}$$

$$\Lambda_2 = \frac{L_2}{L_0} = \rho^{*d_1}\left(\frac{t_1 + t_2}{\rho^* t_1 + t_2}\right)^{d_1+d_2}$$

As in section 2.B. we terminate sampling as soon as $\max(\Lambda_1,\Lambda_2)$ leaves the interval (A,B) where $0 < A < 1 < B < \infty$. This test is due to Flehinger and Louis (10), and is used in conjunction with each allocation rule in the next section.

D. Allocation Rules

1. The R_γ Rules

A class of allocation rules, proposed and examined by Flehinger and Louis (10), are the analogues of the R_γ rules for the normal case described in the previous chapter. As in that case, a constant $0 \leq \gamma \leq 1$ is chosen and kept fixed throughout the experiment. At the time of arrival of the $(n + 1)^{st}$ subject a treatment is assigned to him according to the probability p_{n+1} of assigning treatment 1:

$$p_{n+1} = \begin{cases} 1, & d_2 - d_1 \geq \gamma(n + 1), \text{ or } |d_1 - d_2| < \gamma(n + 1) \text{ and } \hat{\lambda}_1 \leq \hat{\lambda}_2 \\ 0, & d_1 - d_2 \geq \gamma(n + 1), \text{ or } |d_1 - d_2| < \gamma(n + 1) \text{ and } \hat{\lambda}_2 \leq \hat{\lambda}_1 \end{cases}$$

$$p_1 = \frac{1}{2}$$

The effect of this rule is to assign the treatment on which fewer subjects have died, if the difference between death numbers is sufficiently large. If the difference is small, the rule chooses the treatment with the lower estimated death rate.

The indeterminacy in the case $d_1 = d_2 = 0$ is avoided by assigning alternately until at least one death has been observed. Equivalently,

we may base the rule on $\hat{\mu}_i = \lambda_i^{-1}$, taking Bartholomew's (5) version of this estimate in the case $d_i = 0$. That is, if $d_i = 0$, then $\hat{\mu}_i = t_i$, $i = 1,2$.

a. Characteristics of the R_γ rules

Estimated operating characteristic (OC), average sample number (ASN), and inferior treatment number (ITN) surfaces were given by Flehinger and Louis (10), results which we were able to confirm with a reasonable degree of accuracy with our own Monte Carlo trials. Detailed output from these trials revealed several noteworthy features.

(i) The ITN appears to be minimized for some γ near 0.5, for all λ_1 and λ_2.

(ii) For all λ_1 and λ_2, the ASN increases with γ. Analogous to the normal theory results of Robbins and Siegmund, the ASN seems to be minimized by strict alternate sampling. (See Table 3 on page 37.)

(iii) Sampling tends to proceed in long runs of each treatment.

(iv) Let $\rho = \lambda_1/\lambda_2$ be fixed. Then the ratio ITN/ASN decreases in λ_1 for a fixed ρ.

(v) For a given stopping rule, the OC function depends only on ρ. (See Table 4 on page 38.)

In summary, it appears that the R_γ rules oversample on the better treatment, particularly for small values of λ_1 and λ_2. This may be due to the tendency to stick on a single treatment. If the sampling pattern were more mixed, the amount of information on each treatment would increase at all times. The choice of γ presents a considerable problem. Increasing γ from 0.2 to 0.5 gives a small decrease in the

TABLE 3

AVERAGE SAMPLE NUMBERS UNDER VARIOUS

ALLOCATION RULES, FOR SELECTED MEAN LIFETIMES

λ_1^{-1}	λ_2^{-1}	Rule	ASN
10	10	Alternation	75.1
		R_0	77.8
		$R_{.2}$	78.7
		$R_{.5}$	98.0
10	20	Alternation	68.9
		R_0	69.5
		$R_{.2}$	78.1
		$R_{.5}$	108.8
100	200	Alternation	136.3
		R_0	137.4
		$R_{.2}$	183.0
		$R_{.5}$	310.3
100	50	Alternation	111.0
		R_0	110.1
		$R_{.2}$	127.3
		$R_{.5}$	199.7

Termination by SLRT with $A = 0.1$, $B = 30$, $\rho^* = 2$

TABLE 4

OPERATING CHARACTERISTICS OF SELECTED ALLOCATION RULES,

FOR SELECTED MEAN LIFETIMES

ρ	Rule	Probability of Selecting Better Treatment	
		$\lambda_1^{-1} = 10$	$\lambda_2^{-1} = 50$
3	Alternation	1.00	1.00
	$R_{.2}$	1.00	1.00
	$R_{.5}$	1.00	1.00
2	Alternation	0.94	0.98
	$R_{.2}$	0.96	0.91
	$R_{.5}$	0.92	0.91
1	Alternation	(0.95)	(0.98)
	$R_{.2}$	(0.96)	(0.96)
	$R_{.5}$	(0.96)	(0.96)
1/2	Alternation	0.95	0.95
	$R_{.2}$	0.93	0.93
	$R_{.5}$	0.97	0.96
1/3	Alternation	1.00	1.00
	$R_{.2}$	1.00	1.00
	$R_{.5}$	1.00	1.00
1.5	Alternation	0.52	0.47
	$R_{.2}$	0.53	0.62
	$R_{.5}$	0.48	0.58
2/3	Alternation	0.58	0.52
	$R_{.2}$	0.48	0.54
	$R_{.5}$	0.55	0.51

Termination by SLRT with A = 0.1, B = 30, $\rho^* = 2$

ITN at the expense of a (perhaps unacceptable) increase in the ASN.

We now examine two rules which avoid the use of an arbitrary constant.

2. Likelihood-Bias Sampling

Hoel and Weiss (14) used likelihood-bias (L-B) sampling in conjunction with a modification of procedure P_B^* of Bechhofer, Kiefer, and Sobel (6). In the two-population case, this procedure calculates P_1 and P_2 after each observation, where

$$P_i = P[\text{correctly selecting treatment } i \mid \text{observations}], \quad i = 1,2.$$

L-B sampling uses the posterior probabilities P_1 and P_2 to adjust the sampling ratio by defining

$$P(\text{assigning treatment 1}) = P_1/(P_1 + P_2).$$

Note that this is a randomized allocation rule.

Using the SLRT, Sobel and Weiss carried out Monte Carlo trials to compare L-B sampling with the R_γ rules. They found the R_γ rules to be markedly superior in terms of both the ASN and the ITN.

3. Rule R_e

This rule attempts to "mix" the observations better than the R_γ rules, while avoiding the use of an arbitrary constant. Unlike other allocation rules, it does not depend on the observations through the likelihood ratio. Instead, it forces the sampling ratio n_1/n to approach the ratio $\hat{\mu}_1/(\hat{\mu}_1 + \hat{\mu}_2)$. This was motivated by the observation during the initial Monte Carlo trials on R_γ, that these two ratios tended to be approximately equal.

Define $$\alpha_1 = \left| \frac{n_1 + 1}{n + 1} - \frac{\hat{\mu}_1}{\hat{\mu}_1 + \hat{\mu}_2} \right|$$

$$\alpha_2 = \left| \frac{n_1}{n + 1} - \frac{\hat{\mu}_1}{\hat{\mu}_1 + \hat{\mu}_2} \right|$$

Then $p_1 = 1/2$ and $p_{n+1} = 1 - p_n$, if either of d_1 or d_2 is zero, and

$$p_{n+1} = \begin{cases} 1, & \alpha_1 < \alpha_2 \\ 0. & \alpha_1 > \alpha_2 \end{cases},$$

if neither d_1 nor d_2 are zero.

Monte Carlo trials indicated that this rule succeeds well in mixing the observations. However, when the first subject dies almost immediately, $\hat{\mu}_1/(\hat{\mu}_1 + \hat{\mu}_2)$ is close to 0 or 1, and essentially no mixing occurs at all; the procedure may in fact fail to terminate.

The problem is due to the instability of $\hat{\mu}_1/(\hat{\mu}_1 + \hat{\mu}_2)$ for small sample sizes. We can ensure greater stability by alternating until a sufficient number of deaths has been recorded; that is, alternation ceases when

$$n = \min\{n : \max(d_{1n}, d_{2n}) = k\}.$$

k is chosen so that the procedure cannot terminate under alternation, by solving $\Lambda(n) < B$ for k. This gives

$$k = [\log_{\rho^*} B] + 1,$$

the square bracket denoting the greatest integer function.

Monte Carlo results incorporating this "delay" showed that rule R_e gives a lower ASN than the R_γ rules, but the R_γ rules tend to have lower ITN's.

4. Delayed R_γ Rules

Introducing a delay to rule R_e gave a substantial reduction in ASN. It is therefore reasonable to ask if such a delay would similarly improve the R_γ rules. Monte Carlo trials showed that this was the case, along with a reduction in the ITN.

Tables 5, 6, and 7 (shown on pages 42, 43, and 44) give Monte Carlo results for rules R_e, R_0, and $R_{.2}$, all with and without delays. In all cases $\rho^* = 2$, and termination took place with $A = 0.1$, $B = 30$. In delayed trials, the delay used was 5 deaths. The delayed R_γ rules have lower ASN's and ITN's, with relatively little change in their operating characteristics from the non-delayed rules. Rule R_e is competitive, especially for values of ρ greater than 2 or less than 1/2.

E. Summary

It is possible to reduce the inferior treatment number of a two-population sequential test, by using data-dependent treatment allocation. In the case of comparing the parameters of two exponential populations, we have presented two rules, delayed R_e, and delayed R_γ, which improve upon previously published results.

TABLE 5

CHARACTERISTICS OF RULE R_e WITH AND WITHOUT DELAY

λ_2^{-1} = 200		ASN		ITN		OC	
λ_1^{-1}	ρ	Delayed	Not Delayed	Delayed	Not Delayed	Delayed	Not Delayed
600	3	148	169	44	48	1.00	1.00
500	2.5	156	172	48	51	.99	1.00
400	2	198	213	79*	81	.96	.94
300	1.5	223	231	91	100	.48	.54
266.67	1.333	221	223	96	97	.30, .1	.37
200	1	187	182	-	-	.04	.06
150	.75	202	187	88	81	.22	.22
133.33	.667	191	197	77	79	.58	.55
100	.5	138	134	47	45	.96	.92
80	.4	109	114	34*	37	1.00	1.00
66.667	.333	92	94	27	27	1.00	1.00
λ_2^{-1} = 10							
30	3	55	67	15	18	1.00	1.00
25	2.5	60	76	18*	21	1.99	.99
20	2	80	90	27*	31	.96	.92
15	1.5	108	102	43	40	.62	.60
13.333	1.333	106	106	45	46	.35	.36
10	1	80	87	-	-	.02	.07
7.5	.75	90	95	38	41	.29	.33
6.667	.667	97	96	39	44	.51	.56
5	.5	71	77	24	24	.95	.93
5	.4	56	67	16*	18	1.00	.99
3.333	.333	51	55	13	13	1.00	1.00

Termination by SLRT with A = 0.1, B = 30.

100 trials per parameter pair

* significant at $p < 0.05$

TABLE 6

CHARACTERISTICS OF RULE R_0 WITH AND WITHOUT DELAY

λ_2^{-1} = 200		ASN		ITN		OC	
λ_1^{-1}	ρ	Delayed	Not Delayed	Delayed	Not Delayed	Delayed	Not Delayed
600	3	148	155	43*	48	1.000	.999
500	2.5	157	171	51*	58	.995	.993
400	2	194	196	73*	76	.925	.949
300	1.5	226	231	100	102	.525	.573
266.67	1.333	222	222	103	103	.345	.307
200	1	182	185	-	-	.045	.053
150	.75	189	198	90	91	.265, .005	.302
133.33	.667	189	196	83*	86	.580	.570
100	.5	142	146	55*	58	.920	.950
80	.4	113	113	39	41	.995	.994
66.67	.333	92	97	29*	33	1.000	.999
λ_2^{-1} = 10							
30	3	48	47	19	19	1.000	.999
25	2.5	56	55	24	24	.990	.993
20	2	68	71	31	32	.945	.943
15	1.5	100	99	49	48	.565	.572
13.333	1.333	93	97	46	48	.340, .005	.298
10	1	76	76	-	-	.070	.050
7.5	.75	90	94	44	46	.305	.247
6.667	.667	96	96	47	47	.555	.573
5	.5	65	65	30	31	.93	.947
4	.4	45	50	20*	23	.995	.995
3.333	.333	36	45	16	21	.995	.998

Termination by SLRT with A = 0.1, B = 30.

200 trials per parameter pair

* significant at $p < 0.05$

TABLE 7

CHARACTERISTICS OF RULE $R_{0.2}$ WITH AND WITHOUT DELAY

λ_2^{-1} = 200		ASN		ITN		OC	
λ_1^{-1}	ρ	Delayed	Not Delayed	Delayed	Not Delayed	Delayed	Not Delayed
600	3	152	218	33	35	1.000	.998
500	2.5	175	231	36*	43	.995	.994
400	2	212	249	47*	58	.915	.949
300	1.5	265	273	78*	86	.555	.561
266.67	1.333	248	256	77	90	.345	.292
200	1	210	211	-	-	.040	.051
150	.75	221	225	77	81	.255	.284
133.33	.667	219	226	70	74	.530	.551
100	.5	158	178	44	46	.920	.943
80	.4	124	150	26*	33	1,000	.995
66.67	.333	101	134	22	26	1.000	.999
λ_2^{-1} = 10							
30	3	57	56	17	17	1.000	.999
25	2.5	64	63	21	21	1.000	.993
20	2	75	79	26*	28	.935	.953
15	1.5	103	106	40	42	.525	.559
13.333	1.333	98	102	41	42	.300	.303
10	1	77	79	-	-	.035	.048
7.5	.75	90	98	38	41	.295	.297
6.667	.667	94	101	38	41	.550	.537
5	.5	66	73	25	27	.940	.948
4	.4	47	57	17*	21	1.000	.993
3.333	.333	41	53	14	20	1.000	.999

Termination by SLRT with A = 0.1, B = 30.

200 trials per parameter pair

* significant at $p < 0.05$

APPENDIX

COMPUTER PROGRAMS

The programs used to derive the Monte Carlo results of Chapter IV are presented below. Program MC1 simulates the R_γ rules, and MC2 simulates R_e. Both programs call subroutine DATA, which allocates storage for data generated by subroutine GGEXP (from the International Mathematical Subroutines Library).

All procedures are written in FORTRAN IV, for execution on the University of Alberta Computing System.

Following are the input variables for the two programs.

1. NRUNS - The number of stopping parameter sets.
2. ISEED - The "seed" required by the random-number generation subroutines.
3. KRUNS - The number of replications desired under each set of parameters.
4. NALT - The length of the delay.
5. KARDS - The number of pairs of exponential parameters to be considered.
6. SL, SLF, SU, SUF - A stopping parameter set. SL corresponds to A, and SU to B. If SLF and SUF differ from SL and SU, respectively, the test has triangular boundaries.
7. GAMMA - The γ for the R_γ rules.
8. ST - The maximum number of trials to be made. This was usually taken as 1200, a number never exceeded in the simulations.
9. C - The critical ratio (ρ).

```
C              PROGRAM MC1
C                  SIMULATION OF R(GAMMA) WITH TRUNCATION
C
      COMMON ISEED,X(1200,2),E(2),KALL(2),KLM(2),NP
      REAL T(2),R(2),Y
      REAL X,E,EH(2),AM(2)
      INTEGER KALL,KLM,NP
      INTEGER D(2),N(2),MT(2)
      INTEGER ISEED
      READ(7,500) NRUNS,KRUNS,ISEED
      DO 75 KRUNCH=1,NRUNS
      READ(7,501) NALT,KARDS,SL,SLF,SU,SUF,GAMMA,ST,C
      WRITE(6,408) SL,SLF,SU,SUF,ST,C,NALT
      DO 75 L=1,KARDS
      READ(7,502) E(1),E(2)
      WRITE(6,407) E(1),E(2),GAMMA
      IN=2
      IF(E(1).LT.E(2)) IN=1
      PTN=0.
      ASN=0.
      AM(1)=0.
      AM(2)=0.
      KH0=0
      KH1=0
      KH2=0
      A2=0.
      P2=0.
      MAXN=0
C
C                    INDIVIDUAL RUNS START HERE
C
      DO 70 KOUNT=1,KRUNS
      NN=0
      KALL(1)=1
      KALL(2)=1
      KLM(1)=50
      KLM(2)=50
      DO 3 I=1,2
      NP=I
C
C           SINGLE-RUN INITIALIZATION
C
    3 CALL DATA
      DO 4 I=1,2
      T(I)=0.
      D(I)=0
      N(I)=0
    4 R(I)=1.
      NP=1
      CALL GGUB(ISEED,1,Y)
      IF(Y.LT.0.5) NP=2
C
C                    SAMPLING AREA
C
    1 N(NP)=N(NP)+1
      NN=NN+1
```

```
      IF(NN.GT.ST) GO TO 30
      IF(N(NP).LE.KLM(NP)) GO TO 5
      CALL DATA
    5 DO 15 J=1,2
      IF(N(J).EQ.0) GO TO 15
      K=N(J)
      DO 10 I=1,K
      IF(X(I,J).EQ.0.) GO TO 10
      X(I,J)=X(I,J)-1.
      IF(X(I,J).GE.0.) GO TO 9
      D(J)=D(J)+1
      T(J)=T(J)+X(I,J)+1.
      X(I,J)=0.
      GO TO 10
    9 T(J)=T(J)+1.
   10 CONTINUE
      IF((T(1)+T(2)).EQ.0.) GO TO 15
      R(J)=D(3-J)*ALOG(C)+(D(1)+D(2))*ALOG((T(1)+T(2))/
     %(C*T(3-J)+T(J)))
      R(J)=EXP(R(J))
   15 CONTINUE
      M=IABS(D(1)-D(2))
      GO TO 16
   12 NP=3-NP
      GO TO 1
C
C                TERMINATION RULES APPLIED
C
   16 SB=SU+(SUF-SU)*NN/ST
      SA=SL+(SLF-SL)*NN/ST
      AR=AMAX1(R(1),R(2))
      IF((AR.LT.SA).OR.(AR.GT.SB)) GO TO 30
C
C                   ALLOCATION RULES APPLIED
C
C                       R(GAMMA) RULE
C
      IF(MAX0(D(1),D(2)).LT.NALT) GO TO 12
      NP=2
      IF(M.GT.(GAMMA*NN)) GO TO 20
      IF(T(1)+T(2).EQ.0.) GO TO 20
      IF((D(1)/T(1)).LE.(D(2)/T(2))) NP=1
      GO TO 1
   20 IF(D(2).GT.D(1)) NP=1
      GO TO 1
C
C          OUTPUT AREA
C
   30 IF(D(1).EQ.0) EH(1)=1.23456
      IF(D(1).NE.0) EH(1)=T(1)/D(1)
      IF(D(2).EQ.0) EH(2)=1.23456
      IF(D(2).NE.0) EH(2)=T(2)/D(2)
      MT(1)=N(1)-D(1)
      MT(2)=N(2)-D(2)
      MAXN=MAX0(MAXN,NN)
      WRITE(3,414) MT(1),MT(2),N(1),N(2),D(1),D(2),T(1),T(2),
```

```
     &EH(1),EH(2)
       IF(AR.LT.SA) GO TO 40
       IF(NN.GT.ST) GO TO 45
       NP=1
       IF(R(2).GT.R(1)) NP=2
       GO TO (201,202),NP
  201  KH1 = KH1 + 1
       GO TO 50
  202  KH2 = KH2 + 1
       GO TO 50
   40  WRITE(3,401) NN
       KH0 = KH0 + 1
       GO TO 65
   45  AH=ALOG(EH(1)/EH(2))
       IF(ABS(AH).LT.ALOG(C)) GO TO 40
       NP=1
       IF(AH.LT.0.0) NP=2
   50  WRITE(3,402) NN,NP
   65  PTN=PTN+N(IN)
       ASN=ASN+NN
       A2=A2+(FLOAT(NN))**2
       P2=P2+N(IN)**2
       AM(1)=MT(1)+AM(1)
       AM(2)=MT(2)+AM(2)
   70  CONTINUE
C
C          MULTI-RUN SUMMARY
C
       RUNS=FLOAT(KRUNS)
       ASN=ASN/RUNS
       PTN=PTN/RUNS
       A2=(A2-RUNS*(ASN**2))/(RUNS-1.)
       P2=(P2-RUNS*(PTN**2))/(RUNS-1.)
       AM(1)=AM(1)/RUNS
       AM(2)=AM(2)/RUNS
       WRITE(6,405) ASN,PTN,A2,P2,MAXN
       WRITE(6,411) AM(1),AM(2)
    75 WRITE(6,406) RUNS,KH0,KH1,KH2
       STOP
C
C          5XX=INPUT     4XX=OUTPUT
C
  500   FORMAT(I2,I4,I9)
  501   FORMAT(I4,I2,6F6.3,F4.1)
  502   FORMAT(2F6.2)
  407   FORMAT('1  MEAN LIFE ONE = ',F6.2,'  MEAN LIFE TWO = ',F6.2,
     &20X,'GAMMA = ',F4.2)
  401   FORMAT(' AFTER ',I3,' TRIALS NO DIFFERENCE FOUND')
  402   FORMAT(' AFTER ',I3,' TRIALS TREATMENT ',I1,' CHOSEN')
  405   FORMAT('0 ASN = ',F5.1,' ITN = ',F5.1,10X,'VARSN = ',F10.2,
     &'  VARIN = ',F10.2,'       MAXN = ',I4)
  406   FORMAT('  RUNS = ',F4.0,'  H0 ',I4,'  H1 ',I4,'  H2 ',I4)
  408   FORMAT('1',////,'LOWER TERMINATION AT ',F4.3,' TO ',F4.3,
```

```
     &///,'UPPER TERMINATION AT ',F4.1,' TO ',F4.1,///,
     &'TRUNCATION AT ',F6.1,//////,'CRITICAL RATIO ',F4.1,
     &///,'ALTERNATE UNTIL ',I4)
411   FORMAT(' AVERAGE NUMBER OF ACTIVES: ',F6.1,6X,F6.1)
414   FORMAT(1X,6I6,4E14.6)
      END
```

```
C                    PROGRAM MC2
C                  SIMULATION OF R(E) WITH TRUNCATION
C
      COMMON ISEED,X(1200,2),E(2),KALL(2),KLM(2),NP
      REAL T(2),R(2),Y,CH(2)
      REAL X,E,EH(2),AM(2)
      INTEGER KALL,KLM,NP
      INTEGER D(2),N(2),MT(2)
      INTEGER ISEED
      READ(7,500) NRUNS,KRUNS,ISEED
      DO 75 KRUNCH=1,NRUNS
      READ(7,501) NALT,KARDS,SL,SLF,SU,SUF,ST,C
      WRITE(6,408) SL,SLF,SU,SUF,ST,C,NALT
      DO 75 L=1,KARDS
      READ(7,502) E(1),E(2)
      WRITE(6,407) E(1),E(2)
      IN=2
      IF(E(1).LT.E(2)) IN=1
      PTN=0.
      ASN=0.
      AM(1)=0.
      AM(2)=0.
      KH0=0
      KH1=0
      KH2=0
      A2=0.
      P2=0.
      MAXN=0
C
C                  INDIVIDUAL RUNS START HERE
C
      DO 70 KOUNT=1,KRUNS
      NN=0
      KALL(1)=1
      KALL(2)=1
      KLM(1)=50
      KLM(2)=50
      DO 3 I=1,2
      NP=I
C
C          SINGLE-RUN INITIALIZATION
C
    3 CALL DATA
      DO 4 I=1,2
      T(I)=0.
      D(I)=0
      N(I)=0
    4 R(I)=1.
      NP=1
      CALL GGUB(ISEED,1,Y)
      IF(Y.LT.0.5) NP=2
C
C                  SAMPLING AREA
C
    1 N(NP)=N(NP)+1
      NN=NN+1
```

```
      IF(NN.GT.ST) GO TO 30
      IF(N(NP).LE.KLM(NP)) GO TO 5
      CALL DATA
    5 DO 15 J=1,2
      IF(N(J).EQ.0) GO TO 15
      K=N(J)
      DO 10 I=1,K
      IF(X(I,J).EQ.0.) GO TO 10
      X(I,J)=X(I,J)-1.
      IF(X(I,J).GE.0.) GO TO 9
      D(J)=D(J)+1
      T(J)=T(J)+X(I,J)+1.
      X(I,J)=0.
      GO TO 10
    9 T(J)=T(J)+1.
   10 CONTINUE
      IF((T(1)+T(2)).EQ.0.) GO TO 15
      R(J)=D(3-J)*ALOG(C)+(D(1)+D(2))*ALOG((T(1)+T(2))/
     %(C*T(3-J)+T(J)))
      P(J)=EXP(R(J))
   15 CONTINUE
      M=IABS(D(1)-D(2))
      GO TO 16
   12 NP=3-NP
      GO TO 1
C
C                TERMINATION RULES APPLIED
C
   16 SB=SU+(SUF-SU)*NN/ST
      SA=SL+(SLF-SL)*NN/ST
      AR=AMAX1(R(1),R(2))
      IF((AR.LT.SA).OR.(AR.GT.SB)) GO TO 30
C
C                ALLOCATION RULES APPLIED
C
C                     R(E) RULE
C
      IF(MAX0(D(1),D(2)).LT.NALT) GO TO 12
      MD=MIN0(D(1),D(2))
      IF(MD.GT.0) TO TO 19
      IF(MD.EQ.D(1)) NP=1
      IF(MD.EQ.D(2)) NP=2
      GO TO 1
   19 B=(T(1)*D(2))/T(1)*D(2)+T(2)*D(1))
      CH(1)=ABS(B-(N(1)+1.0)/(N(1)+N(2)+1.0))
      CH(2)=ABS(B-N(1)/(N(1)+N(2)+1.0))
      DO 20 I=1,2
   20 IF(AMIN1(CH(1),CH(2)).EQ.CH(I)) NP=I
      GO TO 1
C
C        OUTPUT AREA
C
   30 IF(D(1).EQ.0) EH(1)=1.23456
      IF(D(1).NE.0) EH(1)=T(1)/D(1)
      IF(D(2).EQ.0) EH(2)=1.23456
      IF(D(2).NE.0) EH(2)=T(2)/D(2)
```

```
      MT(1)=N(1)-D(1)
      MT.2)=N(2)-D(2)
      MAXN=MAX0(MAXN,NN)
      WRITE(3,414) MT(1),MT(2),N(1),N(2),D(1),D(2),T(1),T(2),
     &EH(1),EH(2)
      IF(AR.LT.SA) GO TO 40
      IF(NN.GT.ST) GO TO 45
      NP=1
      IF(R(2).GT.R(1)) NP=2
      GO TO (201,202),NP
  201 KH1 = KH1 + 1
      GO TO 50
  202 KH2 = KH2 + 1
      GO TO 50
   40 WRITE(3,401) NN
      KH0 = KH0 + 1
      GO TO 65
   45 AH=ALOG(EH(1)/EH(2))
      IF(ABS(AH).LT.ALOG(C)) GO TO 40
      NP=1
      IF(AH.LT.0.0) NP=2
   50 WRITE(3,402) NN,NP
   65 PTN=PTN+N(IN)
      ASN=ASN+NN
      A2=A2+(FLOAT(NN))**2
      P2=P2+N(IN)**2
      AM(1)=MT(1)+AM(1)
      AM(2)=MT(2)+AM(2)
   70 CONTINUE
C
C        MULTI-RUN SUMMARY
C
      RUNS=FLOAT(KRUNS)
      ASN=ASN/RUNS
      PTN=PTN/RUNS
      A2=(A2-RUNS*(ASN**2))/(RUNS-1.)
      P2=(P2-RUNS*(PTN**2))/(RUNS-1.)
      AM(1)=AM(1)/RUNS
      AM(2)=AM(2)/RUNS
      WRITE(6,405) ASN,PTN,A2,P2,MAXN
      WRITE(6,411) AM(1),AM(2)
   75 WRITE(6,406) RUNS,KH0,KH1,KH2
      STOP
C
C        5XX=INPUT     4XX=OUTPUT
C
  500 FORMAT(I2,I4,I9)
  501 FORMAT(I4,I2,6F6.3,F4.1)
  502 FORMAT(2F6.2)
  407 FORMAT('1  MEAN LIFE ONE = ',F6.2,'  MEAN LIFE TWO = ',F6.2)
  401 FORMAT(' AFTER ',I3,' TRIALS NO DIFFERENCE FOUND')
  402 FORMAT(' AFTER ',I3,' TRIALS TREATMENT ',I1,' CHOSEN')
  405 FORMAT('0 ASN = ',F5.1,' ITN = ',F5.1,10X,'VARSN = ',F10.2,
```

```
     &'  VARIN = ',F10.2,'       MAXN = ',I4)
406   FORMAT('  RUNS = ',F4.0,'  H0 ',I4,'  H1 ',I4,'  H2 ',I4)
408   FORMAT('1',////,'LOWER TERMINATION AT ',F4.3,' TO ',F4.3,
     &///,'UPPER TERMINATION AT ',F4.1,' TO ',F4.1,///,
     &'TRUNCATION AT ',F6.1,//////,'CRITICAL RATIO ',F4.1,
     &///,'ALTERNATE UNTIL ',I4)
411   FORMAT(' AVERAGE NUMBER OF ACTIVES: ',F6.1,6X,F6.1)
414   FORMAT(1X,6I6,4E14.6)
      END
```

```
      SUBROUTINE DATA
C     SUBROUTINE FOR GENERATION AND ALLOCATION OF DATA
      COMMON ISEED,Y(1200,2),E(2),KALL(2),KLM(2),NP
      REAL Y,E,YV(50)
      INTEGER KALL,KLM,NMP,ISEED
      CALL GGEXP(ISEED,E(NP),50,YV)
      NNN=50*(KALL(NP)-1)+1
      NNNN=NNN+49
      DO 1000 I=NNN,NNNN
      J=MOD(I,50)
      IF(J.EQ.0) J=50
1000  Y(I,NP)=YV(J)
      KLM(NP)=50*KALL(NP)
      KALL(NP)=KALL(NP)+1
      RETURN
      END
```

BIBLIOGRAPHY

1. T.W. Anderson, "A modification of the sequential probability ratio test to reduce the sample size," Ann. Math. Stat., Vol. 31 (1960), pp. 165-197.

2. F.J. Anscombe, "Sequential medical trials," J.A.S.A., Vol. 58 (1963), pp. 365-383.

3. P. Armitage, "Some developments in the theory and practice of sequential medical trials," Proceedings of the 5th Berkeley Symposium, IV (1967), pp. 791-804.

4. P. Armitage, Sequential Medical Trials, Oxford: Blackwell Scientific Publications (1960).

5. D.J. Bartholomew, "A problem in life testing," J.A.S.A., Vol. 52 (1957), pp. 350-355.

6. R. Bechhofer, J. Kiefer, and M. Sobel, Sequential Identification and Ranking Procedures, Chicago: University of Chicago Press (1968).

7. N. Breslow, "On large sample sequential analysis with applications to survivorship data," Journal of Applied Probability, Vol. 6 (1969), pp. 261-274.

8. N. Breslow and C. Haug, "Sequential comparison of exponential survival curves," J.A.S.A., Vol. 67 (1972), pp. 691-697.

9. B.J. Flehinger and T.A. Louis, "Sequential medical trials with data dependent treatment allocation," Proceedings of the 6th Berkeley Symposium, IV (1972), pp. 43-52.

10. B.J. Flehinger and T.A. Louis, "Sequential treatment allocation in clinical trials," Biometrika, Vol. 58 (1971), pp. 419-426.

11. B.J. Flehinger, T.A. Louis, H. Robbins, and B.H. Singer, "Reducing the number of inferior treatments in clinical trials," Proc. Nat. Acad. Sci. U.S.A., Vol. 69 (1972), pp. 2993-2994.

12. B.K. Ghosh, Sequential Tests of Statistical Hypotheses, Reading: Addison-Wesley (1970).

13. D. Hoel and M. Sobel, "Comparisons of sequential procedures for selecting the best binomial populations," Proceedings of the 6th Berkeley Symposium, IV (1972), pp. 53-69.

14. D. Hoel and G. Weiss, "Comparison of methods for choosing the better of two negative exponential lifetime distributions," in Reliability and Biometry, F. Proschan & R. Serfling, eds., Philadelphia, S.I.A.M. (1974).

15. T.A. Louis, "Two Population Hypothesis Testing with Data-Dependent Allocation," Ph.D. Dissertation, Columbia University (1972).

16. H. Robbins and D.O. Siegmund, "Sequential tests involving two populations," J.A.S.A., Vol. 69 (1974), pp. 132-139.

17. M. Sobel and G. Weiss, "Play-the-winner sampling for selecting the better of two binomial populations," Biometrika, Vol. 57 (1970), pp. 357-65.

18. M. Sobel and G. Weiss, "Recent results on using the play-the-winner sampling rule with binomial selection problems," Proc. 6th Berk. Symp., I (1972), pp. 717-736.

19. A. Wald, Sequential Analysis, New York: John Wiley & Sons (1947).

20. M. Weinstein, "Allocation of subjects in medical experiments," New England Journal of Medicine, Vol. 291 (1974), pp. 1278-1285.

21. M. Zelen, "Play-the-Winner and the controlled clinical trial," J.A.S.A., Vol. 64 (1969), pp. 131-146.

22. M. Zelen, "Application of exponential models to problems in cancer research," Jour. Roy. Stat. Soc., Series A, Vol. 129 (1966), pp. 368-390.